

Using Holistic Grading to Evaluate Writing in Engineering Classes

Ronald L. Miller, Barbara M. Olds
Colorado School of Mines

Summary

This paper describes the results of a classroom experiment in which holistic grading was used to evaluate engineering student writing. We first review the concept of holistic grading as it has been developed by professional writing instructors, and then describe the scoring scheme we developed and how we applied it in class. We also discuss the advantages and disadvantages of holistic writing and share feedback from students who participated in our experiment.

Background

Many engineering faculty would like to include more writing assignments in their classes to help improve the communication skills of their students. The ABET 2000 criteria, for example, explicitly require that "engineering programs demonstrate that their graduates have an ability to communicate effectively" while data from employer surveys consistently mention the importance of good communication skills as a requirement for professional success.^{1,2} The engineering education community also knows from many years of experience that one freshman composition course followed by a few junior-level lab reports and one senior design report is not an adequate amount of writing to develop competence in our students. Much like other cognitive skills that we are trying to engender in our students, improving writing requires practice using a variety of academic tasks and contexts. In an informal survey at the Colorado School of Mines, most faculty members noted that they wanted to include more writing in their courses but had several concerns, chiefly the increased grading burden and their perceived lack of expertise in evaluating written work.

The purpose of this paper is to describe the use of holistic grading as a method to help engineering educators rapidly but reliably evaluate student writing in their courses. We will first review the concept of holistic grading as developed by professional writing teachers and then describe one course in which these techniques have been piloted and assessed, including feedback from students in the course.

Holistic Grading of Writing Assignments

When engineering educators think about grading student writing, most envision a tedious and time-consuming exercise of marking all the grammatical, mechanical, and perhaps stylistic errors they can identify. If they're able, they may also try to assess issues of format, voice, syntax and logic in the text. For the untrained instructor, this can be a difficult, if not impossible, task tantamount to providing an editing service which does little to improve student writing. Professional writing instructors deal with the same issues and have developed a new series of grading strategies collectively known as holistic grading. As described by Cooper³,

Holistic evaluation of writing is a guided procedure for sorting or ranking written pieces. The rater takes a piece of writing and either (1) matches it with another piece in a graded series of pieces or (2) scores it for the prominence of certain features important to that kind of writing or (3) assigns it a letter or number grade. The placing, scoring, or grading occurs quickly, impressionistically, after the rater has practiced the procedure with other raters. The rater does not make corrections or revisions in the paper. Holistic evaluation is usually guided by a holistic scoring guide which describes each feature and identifies high, middle, and low quality levels for each feature.

The advantages of holistic grading have been articulated by Lindemann.⁴ Holistic evaluation assumes each written piece communicates a complete message to a desired audience and therefore should be graded on the overall quality of communication, much as a manager might evaluate an engineer's communication skills based on the overall quality of written documents. As a result, holistic grading allows rapid but useful feedback because scoring guides (termed "rubrics") can be tailored to the specific objectives of each assignment, class, or curriculum. Finally, holistic grading removes much of the subjectivity from grading written work by providing students with specific criteria for each grade position on the scoring rubric.

Holistic grading was originally used to rapidly score and rank essays prepared by students taking national placement exams administered by the Educational Testing Service and others. Collegiate general education and writing programs also employ holistic grading as a method to collect summative evaluation data about the overall writing ability of groups of students.⁴ More recently, Cooper³ has proposed using holistic grading as a way to provide formative feedback to students completing writing assignments in individual courses.

The issue of grading reliability using holistic methods had been studied extensively.³ Many studies have demonstrated reliability values exceeding 90% (considered high enough for summative program evaluation or individual student feedback) if raters: 1) work together to develop an appropriate scoring rubric, and 2) calibrate themselves using sample writing pieces similar to those that will be graded in class. Based on these findings, we decided to experiment with using holistic grading of written work in a class at the Colorado School of Mines (CSM) involving multiple instructors.

An Application

Our first application of holistic grading occurred in an interdisciplinary course on technology and society taught yearly by one of the authors as part of the McBride Honors Program in Public Affairs for Engineers at CSM, a 24 credit sequence of seminars and a practicum for students who wish to "explore the interfaces between their areas of technical expertise and the humanities and social sciences; to gain the sensitivity to project and test the moral and social implications of their future professional judgments and activities; and to foster their leadership abilities in preparation for managing change and promoting the general welfare in an evolving technological and global context."⁵ Honors students are expected to achieve several goals including:

- ⇒ the ability to communicate effectively orally and in writing to a variety of audiences
- ⇒ the ability to analyze and critically evaluate both their own ideas and those of others
- ⇒ development of a reflective mind
- ⇒ love of learning and the promise of continuing learning throughout their lives

"Technology and Socio-Economic Change" (TSEC) is a junior-level seminar moderated by CSM engineering and science faculty designed to critically analyze the impact of science and technology on American values and institutions. The role of technology in American society is studied and the implications of technology transfer from developed to developing nations are debated. As an outcome of the TSEC seminar, students learn to relate technological growth to socio-economic, cultural, and religious aspects of society and probe into the moral and social consequences of technological innovations. Seminar sessions consist primarily of full group discussions, small group discussions, and work on a semester-long group (3-4 student) term project on a topic related to one or more seminar themes. Each group's specific term project topic is developed with guidance and input from one of the faculty moderators who then acts as advisor to the group as the term project is completed. As in all honors seminars, students in the TSEC seminar write extensively; assignments such as weekly reading critiques, term project reports, and essay exam questions emphasize argumentation and critical analysis rather than summary and opinion.

Because the TSEC seminar involves many written assignments (approximately 15 per student per semester) and since the seminar moderators are predominately engineering and science faculty rather than trained writing instructors, we recently implemented the use of holistic grading in the seminar. The purpose of this experiment was to find out if we could provide students with useful feedback on their writing (especially their weekly critiques of the assigned seminar readings) without the need for extensive editing of each written assignment. The reading critiques were generally 2-3 pages long; in each paper, students were asked to briefly summarize the readings assigned for each seminar and then critically analyze one or two of the major arguments developed by the author(s).

The first step in this process was developing the scoring rubric shown in Table I [adapted from ref. 6]. We specifically chose a scoring scale different from 1-10 or 1-100 to avoid the implication of percentage scores that students automatically relate to traditional grades.

The rubric was provided in written form to students at the first seminar and briefly discussed. At that time, we noted that we would generally not mark detailed editorial comments on their papers but would be happy to discuss weaknesses in their writing if desired. We also emphasized the need to provide a well-argued critique rather than simply a summary of each class reading and that the quality of each paper's ideas and writing would be judged holistically (that is, separate grades would not be given for the critique and for writing quality). We hoped the scoring rubric would help students understand our expectations for their writing by making clear the requirements for a top grade of 6.

Table I - Holistic Scoring Rubric

- 6 This score will be used for a superior paper which addresses itself to all aspects of the reading. Though it may have occasional faults, the paper will be well organized, detailed, and well written.
 - 5-4 These scores will be used for a well-handled paper which is weak in some aspect compared with a superior paper; for example, it may slight one or more major issues from the reading, it may not be as clearly organized as a superior paper, or it may have some minor grammatical inconsistencies.
 - 3 This score will be used for papers which are only summative, those which do not fully develop a critique, those which fail to provide evidence to support the critique, or those which are general and superficial.
 - 2 This score will be used for papers which exhibit serious weaknesses in structure, syntax, diction, and/or idea development.
 - 1 This score will be used for papers which are non-responsive to the assignment.
-

Our next major task was to ensure uniformity of grading using the scoring rubric (a process called "norming"). Each student's first reading critique was graded independently by at least two of the five seminar moderators. In addition, two moderators independently graded all 55 student papers. We were pleased and somewhat surprised with the results of this exercise -- nearly all papers were independently judged to have the same rubric score. Only about 8 papers received scores which differed by 1 point and only one paper received scores which differed by 2 points. Although this result is commonly observed³, we were happy that we could expect consistent grading among moderators with little formal training. Beginning with the second set of student papers, only one moderator scored each paper.

During the semester, we saw significant improvement in students' writing as well as the ability to critically analyze and argue an issue using evidence. Critique scores during the first few seminars clustered in the 3-4 range; students were able to adequately summarize the readings, but were not capable of criticizing what they read beyond superficial phrases such as "I liked this reading" or "I didn't like the author's style." Even though they received few editorial comments, nearly all students improved steadily and were consistently receiving 5's or 6's on their papers by about the fifth seminar. We noted that the quality of students' writing (grammar, mechanics, style, organization, etc.) and the quality of argumentation improved simultaneously.

Once the moderators became comfortable scoring papers holistically, an individual two page critique could be scored reliably in a few minutes. Some moderators continued to provide written comments on papers, but others simply provided a holistic score with little additional feedback.

Feedback from Students

Since holistic grading of student writing was new to the TSEC seminar, we asked the students to give us their perceptions about the utility of our grading strategy and if they believed their writing had improved during the seminar. As shown in Table II, the perception questionnaire consisted of five questions to which students responded using a 5-point Likert scale; in addition, students were asked to list two positive aspects and two negative aspects of holistic grading. The questionnaire was administered during the last seminar so that students could give us feedback on the entire semester of holistic grading.

Overall, student feedback was mixed. For example, as shown in Figure 1, approximately 40% of the students believed that they received useful feedback on their papers. However, nearly 49% disagreed or strongly disagreed that holistic feedback was useful to them. Only about 30% of the students believed their writing had improved during the semester (Figure 2), despite a steady increase in scores on weekly written assignments and faculty observations that nearly all students' writing significantly improved. The issue of subjectivity in holistic grading also resulted in a wide range of responses (Figure 3). About 39% of the students disagreed or strongly disagreed that holistic grading was too subjective, but 33% agreed or strongly agreed that it was.

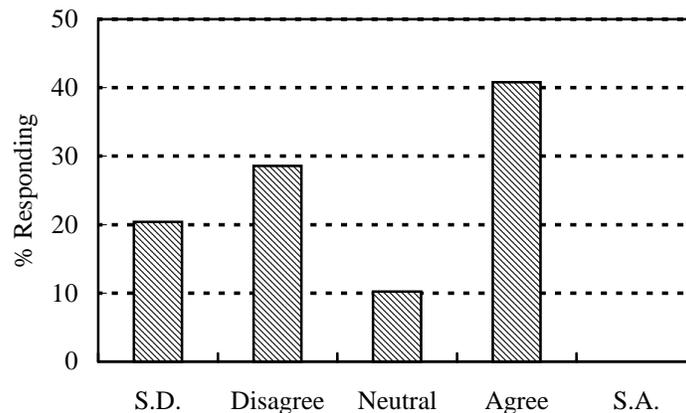


Figure 1 - Student Responses to Questionnaire Statement “Holistic grading provides me with useful feedback on my written papers.” (S.D. = strongly disagree, S.A. = strongly agree)

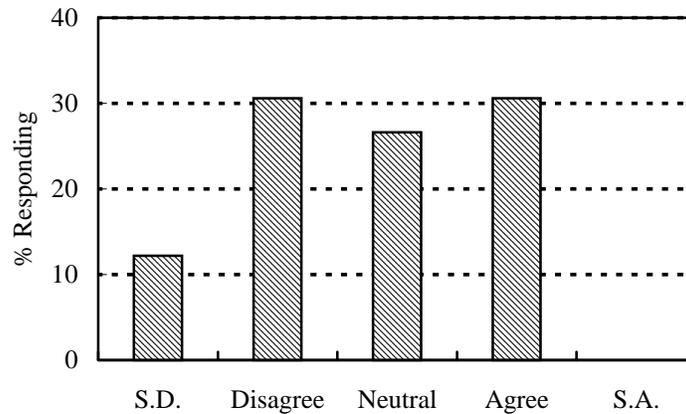


Figure 2 - Student Responses to Questionnaire Statement “My writing has improved this semester.” (S.D. = strongly disagree, S.A. = strongly agree)

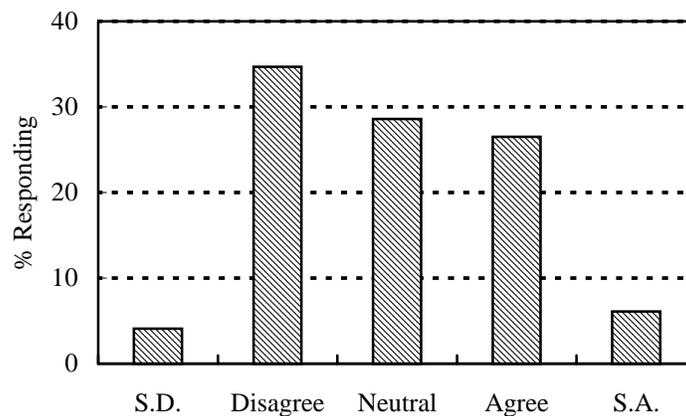


Figure 3 - Student Responses to Questionnaire Statement “Holistic grades seems too subjective.” (S.D. = strongly disagree, S.A. = strongly agree)

Table III summarizes student comments describing their views about positive and negative aspects of holistic grading. Once again, responses were mixed, especially regarding the issues of meaningful feedback and grading subjectivity. Some students commented favorably about the quick evaluation of their papers (about two days from the time papers were emailed to moderators), the focus on overall quality of each paper (writing and content), lack of grade reductions for "nit-picking" errors, and well defined grading criteria provided by the scoring rubric shown in Table I. However, other students commented on the lack of detailed written comments on their papers, perceived subjectivity in grading, and confusion about how holistic grades on a 1-6 scale would be included in each student's course grade.

Clearly, the students were not entirely sold on holistic grading, even though we believe that nearly all of them benefited from the feedback they received about their writing. In analyzing the results summarized above and after informally talking to some of the students in the seminar, we identified several issues which must be resolved before students accept holistic grading as a viable substitute to traditional grading methods. First, students were apparently uneasy with the concept of formative evaluation -- the idea that they were receiving feedback via the scoring rubric that would help them improve their papers, but were not necessarily receiving part of their course grade. Several times during the semester, we told the students that their grade for the reading critiques would be based on the overall quality and improvement in their papers and would not simply be computed as a normed average of their holistic scores. To us, this seemed like an advantage, since students would not have to worry about a few low scores early in the semester and could concentrate on preparing high quality papers by the end of the semester. Instead, feedback indicated that students were uncomfortable that they didn't know their critique grade (and therefore, their weighted class average) at all times during the semester.

Several students commented that a 0-100 scale was more objective than a 1-6 scale, even though experts on holistic grading methods indicate just the opposite.^{3,4} Once again, we believe students were confusing formative feedback (which can be provided using any well-defined scale) with grades based on percentages. Student uneasiness about their perceived lack of objectivity in the scoring scale extended to comments about differences in scores given by each moderator, even though we discussed with them the positive results of our "norming" process. Perhaps variations in the number of written comments provided by each moderator contributed to this perception.

Conclusions and Recommendations

The results of this experiment indicate that holistic grading has potential as a method for providing rapid, reliable, and useful formative and summative feedback to help students improve their writing skills. The method can be adapted to a variety of courses by tailoring rubrics to meet the objectives of each writing assignment. However, several aspects of holistic grading, including grading scales which are not directly related to grade percentages, an absence of detailed editorial comments, and the purpose of formative evaluation are unfamiliar concepts to engineering students. These issues must be addressed explicitly throughout the semester to help students understand the value of holistic grading strategies.

References Cited

- [1] Accreditation Board for Engineering and Technology, "Engineering Criteria 2000," available online at: <http://www.abet.ba.md.us/EAC/eac2000.html>, December 1996.
- [2] "Real World 101: What Some Engineers in Industry Want Your Students, and You, to Know," ASEE Prism, vol. 2, no. 2, pp. 19-22, 1992.
- [3] Cooper, C.R., "Holistic Evaluation of Writing," in Evaluating Writing, C.R. Cooper and L. Odell, eds., National Councils of Teachers of English, Urbana, Illinois, pp. 3-31, 1977.
- [4] Lindemann, E., A Rhetoric for Writing Teachers, Oxford University Press, New York, pp. 214-217, 1982.
- [5] Olds, B.M., and R.L. Miller, "A Liberal Education Model of Leadership Preparation: The McBride Honors Program in Public Affairs for Engineers," Proceedings of the Frontiers in Education Conference (electronic), Salt Lake City, Utah, November 6-9, 1996.
- [6] Moss, A., and C. Holder, Improving Student Writing, Kendall/Hunt Publishing Company, Dubuque, Iowa, pp. 43-46, 1988.

Biographical Information

RONALD L. MILLER is an associate professor of chemical engineering at the Colorado School of Mines. He has been involved in developing and sustaining a number of engineering programs which incorporate writing including EPICS, HumEn, Multidisciplinary Senior Design, the McBride Honors Program, and Connections. He is currently working with a group of faculty to develop a center for educational research at CSM.

BARBARA M. OLDS is a professor of liberal arts and international studies and principal tutor of the Guy T. McBride Honors Program at the Colorado School of Mines. She has been involved in developing and sustaining a number of engineering programs which incorporate writing including EPICS, HumEn, Multidisciplinary Senior Design, the McBride Honors Program, and Connections. She is currently a member of CSM's Curriculum Revision Steering Committee.

Table II
Holistic Grading Questionnaire Used in
"Technology and Socio-Economic Change" Course

This semester we have used holistic grading to evaluate your written work in this course (particularly the reading critiques). To help us improve this grading method, we are interested in your response to the following statements using the scale defined below:

----- ----- ----- -----				
1	2	3	4	5
strongly disagree	disagree	neutral	agree	strongly agree

Remember, there are no "right" or "wrong" answers in a questionnaire like this one; just give us your honest thoughts.

- _____ 1. Holistic grading provides me with useful feedback on my written papers.
- _____ 2. My writing has improved this semester.
- _____ 3. I would prefer receiving grades on a 100 point scale rather than a holistic grading scale.
- _____ 4. Holistic grades seem too subjective.
- _____ 5. Since holistic grading facilitates rapid and useful feedback, it should become the standard method for grading written work in the McBride Honors Program.

List 2 positive aspects of holistic grading:

List 2 negative aspects of holistic grading:

Table III
Summary of Student Responses from Questionnaire

Positive Aspects of Holistic Grading

- quick turnaround
- focus on general quality of writing
- feedback covers more important aspects
- easy to use
- overall quality is reviewed rather than specific points
- effective scale that seems to equalize grading
- have a better idea of what is expected
- grading scale is more standard
- easy to get a good grade
- negates a lot of subjectivity in grading
- easy to interpret score
- graded against known criteria with well distinguished levels
- provides beneficial feedback on my writing
- more incentive for discussion about grades
- rapid and reasonable feedback
- grading is clearly defined
- reduces grading aspects that might be seen as nit-picking
- more standardized grading scheme

Negative Aspects of Holistic Grading

- difficult to tell exactly where you stand gradewise
- personal feelings of profs. come into play
- too subjective
- no subtle differences allowed
- not quantitative enough for engineers
- lower grades

- scale was not large enough -- with 1-100 scale, more room for error
- not used to seeing such low numbers on papers
- no room for unique papers
- grades vary among moderators
- few comments about style, etc.
- difficult to predict grade in course
- very inaccurate to use round numbers such as 1, 2, 3, etc.(needed decimal fractions)
- a "5" is considered good by instructors, but isn't that great
- need more comments on writing style
- not enough justification for the grade given
- papers seemed to be compared to each other
- it's a touchy-feely method